TITLE OF INVENTION

RECOMBINANT *HAEMOPHILUS INFLUENZAE* ADHESIN PROTEINS

REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of copending United States Patent Application No. 09/268,347.

FIELD OF INVENTION

The present invention relates to the field of molecular genetics and, in particular, to the production of recombinant *Haemophilus influenzae* adhesin (Hia) proteins.

BACKGROUND TO THE INVENTION

Haemophilus influenzae is the cause of several serious human diseases, such as meningitis, epiglottitis, septicemia and otitis media. There are six serotypes of *H. influenzae*, designated a to f, that are identified by their capsular polysaccharide. *H. influenzae* type b (Hib) was a major cause of bacterial meningitis until the introduction of several Hib conjugate vaccines in the 1980's (ref. 1. Throughout this application, various references are referred to in parenthesis to more fully describe the state of the art to which this invention pertains. Full bibliographic information for each citation is found at the end of the specification, immediately preceding the claims. The disclosures of these references are hereby incorporated by reference into the present disclosure). Vaccines based upon *H. influenzae* type b capsular polysaccharide conjugated to diphtheria toxoid (ref. 2), tetanus toxoid (ref. 3 and US patent 4,496,538), or *Neisseria meningitidis* outer membrane protein (ref. 4) have been effective in reducing *H. influenzae* type b-

induced meningitis. The other serotypes of *H. influenzae* are associated with invasive disease at low frequencies, although there appears to be an increase in the incidence in disease caused by these strains as the incidence of Hib disease declines (ref. 5; ref. 6). Non-encapsulated or non-typeable *H. influenzae* (NTHi) are also responsible for a wide range of human diseases including otitis media, epiglottitis, pneumonia, and tracheobronchitis. The incidence of NTHi-induced disease has not been affected by the introduction of the Hib vaccines (ref. 7).

Otitis media is the most common illness of early childhood, with 60 to 70% of all children, of less than 2 years of age, experiencing between one and three ear infections (ref. 8). Chronic otitis media is responsible for hearing, speech and cognitive impairments in children. *H. influenzae* infections account for about 30% of the cases of acute otitis media and about 60% of chronic otitis media. In the United States alone, treatment of otitis media costs between 1 and 2 billion dollars per year for antibiotics and surgical procedures such as tonsillectomies, adenoidectomies and insertion of tympanostomy tubes. It is estimated that an additional $30 billion is spent per annum on adjunct therapies, such as speech therapy and special education classes. Furthermore, many of the causative organisms of otitis media are becoming resistant to antibiotic treatment. An effective prophylactic vaccine against otitis media is thus desirable.

During natural infection by NTHi, surface-exposed outer membrane proteins that stimulate an antibody response are potentially important targets for bactericidal and/or protective antibodies and, therefore, potential vaccine candidates. A family of high molecular weight proteins (HMW1 and HMW2) that are important in attachment of NTHi to epithelial cells has been identified in about 70 to 75% of NTHi strains (ref. 9; ref. 10). These high molecular weight adhesins have been shown to afford some protection in the chinchilla model of otitis media (ref. 11). A second family of high molecular weight adhesion proteins has been identified in about 25% of NTHi and in encapsulated *H. influenzae* strains (ref. 12; ref. 13, ref. 14). The NTHi member of this second family is termed *Haemophilus influenzae* adhesin or Hia and the homologous protein found in encapsulated strains is termed *Haemophilus influenzae* surface fibril protein or Hsf. The *hia* gene was originally cloned from an expression library using convalescent sera from an otitis media patient, which indicates that it is an important immunogen during disease. The prototype Hia and Hsf proteins demonstrate about 82% sequence similarity, although the Hsf protein is considerably larger. The proteins are comprised of conserved amino and carboxy termini and several repeat motifs, with Hsf containing more repeat sequences than Hia. A high molecular weight protein (200 kDa) has also been identified from *Moraxella catarrhalis* that has some sequence homology with the Hsf and Hia proteins (U.S. Patent No. 5,808,024).

Since Hia or Hsf is conserved amongst encapsulated strains of *Haemophilus influenzae* and about 20 to 25% of non-encapsulated strains, and has been demonstrated to be an adhesin, the protein has utility in diagnosis of and vaccination against disease caused by *H. influenzae* or other bacterial pathogens that produce Hia or a protein capable of raising antibodies specifically reactive with Hia.

A disadvantage of Hia for use as an antigen in diagnosis, for the generation of anti-Hia antibodies useful in diagnosis and as an immunogen in vaccination is the low recovery of the native protein from *Haemophilus influenzae* species.

It would be advantageous to provide recombinant Hia protein for use as antigens, in immunogenic preparations including vaccines, carriers for other immunogens and in the generation of diagnostic reagents.

SUMMARY OF THE INVENTION

The present invention is directed towards the provision of recombinant *H. influenzae* adhesin (rHia) proteins.

In connection with the provision of such recombinant proteins, the present invention provides certain isolated and purified nucleic acid molecules. Accordingly, in one aspect thereof, the present invention provides an isolated and purified nucleic acid molecule encoding a *Haemophilus influenzae* adhesin (Hia) protein of a strain of *Haemophilus influenzae* having: (a) a DNA sequence selected from the group consisting of those shown in Figures 18, 19, 20, 21,

22, 23, 24 and 25 (SEQ ID Nos: 23, 25, 27, 29, 31, 33, 35, 37); or (b) a DNA sequence encoding a *Haemophilus influenzae* adhesin (Hia) protein having an amino acid sequence selected from the group consisting of those shown in Figures 18, 19, 20, 21, 22, 23, 24 and 25 (SEQ ID Nos: 24, 26, 28, 30, 32, 34, 36, 38).

Such nucleic acid may be included in a vector, which may be a plasmid vector. In particular, the nucleic acid molecule may encode the Hia protein from strain 11 or 33 of non-typeable *Haemophilus*.

In another aspect of the present invention, there is provided an isolated and purified nucleic acid molecule encoding an N-truncated *Haemophilus influenzae* adhesin (Hia) protein of a strain of *Haemophilus influenzae* which is amplifiable by a pair of nucleotides which are selected from the group consisting of SEQ ID No: 7 and SEQ ID No: 15; SEQ ID No: 9 and SEQ ID No: 15; SEQ ID No: 11 and SEQ ID No: 15; SEQ ID No: 13; SEQ ID No: 15; SEQ ID No: 49; and SEQ ID No: 51.

Such nucleic acid may be included in a vector, which may be a plasmid vector. In particular, the nucleic acid molecule may encode an N-truncated Hia protein from strain 11 or 33 of non-typeable *Haemophilus*, starting at codon V38 or S44.

The plasmid vector incorporating the isolated and purified nucleic acid provided in accordance with these aspects of the invention may have the identifying characteristics of a plasmid which is selected from the group consisting of:

DS-2008-2-3 as shown in Figure 1A

DS-2186-1-1 as shown in Figure 5A

DS-2201-1 as shown in Figure 5A

DS-2186-2-1 as shown in Figure 5A

DS-2168-2-6 as shown in Figure 5A

1A-191-3-1 as shown in Figure 32

The vector provided herein may include the *cer* gene from *E. coli*. Accordingly, in another aspect of the present invention, there is provided a vector for transforming a host, comprising a nucleic acid molecule encoding a full-length or N-truncated *Haemophilus influenzae* adhesin (Hia) protein, a promoter for expression of said full-length or truncated Hia protein and, optionally, the *cer* gene of *E. coli*. The vector may be a plasmid vector or other non-replicating vector, which may have the identifying characteristics of a plasmid vector which is selected from the group consisting of:

BK-96-2-11 as shown in Figure 6A

DS-2242-1 as shown in Figure 7A

DS-2242-2 as shown in Figure 7A

DS-2340-2-3 as shown in Figure 8A

DS-2447-2 as shown in Figure 9A

DS-2448-17 as shown in Figure 9B

JB-2930-3 as shown in Figure 32

The vectors provided herein may comprise a replicating vector, including a vector from *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus.

Any of the vectors provided herein may be employed to transform a suitable host cell for expression therein of a protective *Haemophilus influenzae* adhesin (Hia) protein of a non-typeable strain of *Haemophilus*,

which may be in full-length or truncated form. Such host conveniently may be *E. coli*. Such expression may be under the control of the T7 promoter and expression of the recombinant Hia from the transformed host may be effected by culturing in an inducing concentration of lactose or other convenient inducing agent.

The present invention further includes, in a further aspect thereof, a recombinant protective *Haemophilus influenzae* adhesin (Hia) protein of a non-typeable *Haemophilus* strain producible by the transformed host, particularly *E. coli*, provided herein. Such Hia protein may be provided in the form of an immunogenic fragment or adhesin-functional analog of the recombinant protein.

The recombinant Hia proteins, full-length or N-truncated, provided herein are useful as antigens in immunogenic compositions, carriers for other immunogens, diagnostic agents and in the generation of diagnostic agents. The nucleic acid molecules which encode the Hia protein, full-length or N-truncated, also are useful as probes for diagnostic use and also in immunogenic compositions.

The present invention, in an additional aspect thereof, provides an immunogenic composition, comprising at least one immunologically active component which is selected from the group consisting of an isolated and purified nucleic acid molecule as provided herein and a recombinant protective Hia protein, full-length or N-truncated, of a strain of *Haemophilus*, as provided herein, and a pharmaceutically-acceptable carrier therefor.

The immunogenic compositions provided herein may be formulated as a vaccine for *in vivo* administration to a host to provide protection against disease caused by *H. influenzae*. For such purpose, the compositions may be formulated as a microparticle, capsule, ISCOM or liposome preparation. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces.

The immunogenic compositions of the invention (including vaccines) may further comprise at least one other immunogenic or immunostimulating material and the immunostimulating material may be at least one adjuvant or at least one cytokine. Suitable adjuvants for use in the present invention include (but are not limited to) aluminum phosphate, aluminum hydroxide, QS21, Quil A, derivatives and components thereof, ISCOM matrix, calcium phosphate, calcium hydroxide, zinc hydroxide, a glycolipid analog, an octadecyl ester of an amino acid, a muramyl dipeptide, polyphosphazene, ISCOPREP, DC-chol, DDBA and a lipoprotein and other adjuvants.

Advantageous combinations of adjuvants are described in copending United States Patent Application Serial No. 08/261,194 filed June 16, 1994 and 08/483,856 filed June 7, 1995, assigned to the assignee hereof and the disclosure of which is incorporated herein by reference (WO 95/34308 published November 21, 1995).

In accordance with another aspect of the invention, there is provided a method for generating an immune response in a host, comprising the step of

administering to a susceptible host an effective amount of the immunogenic composition as recited above. The immune response may be humoral or a cell-mediated immune response. Hosts in which protection against disease may be conferred include primates, including humans.

In accordance with other aspects of the invention, there is provided the immunogenic compositions provided herein when used as a medicament and the use of these components of the immunogenic compositions in the manufacture of an immunogenic composition.

The present invention includes, in a yet additional aspect thereof, a method for the production of a protective *Haemophilus influenzae* adhesin (Hia) protein of a non-typeable strain of *Haemophilus influenzae*, which comprises:

transforming a host, such as *E. coli*, with a vector comprising a nucleic acid molecule encoding an N-truncated form of the *Haemophilus influenzae* adhesin protein as provided herein,

growing the host to express the encoded truncated Hia, and

isolating and purifying the expressed Hia protein.

The encoded truncated Hia may be expressed in inclusion bodies. The isolation and purification step may be effected by disrupting the grown transformed cells to produce a supernatant and the inclusion bodies containing the Hia, solubilizing the inclusion bodies after separation from the supernatant, to produce a solution of the recombinant Hia, chromatographically purifying the solution of recombinant Hia free from

cell debris, and isolating the purified recombinant Hia protein.

The vector transforming the host cell, such as *E. coli*, may include the T7 promoter and the *E. coli* or other host cell may be cultured in the presence of an inducing amount of lactose or other convenient inducing agent.

The strain of *Haemophilus influenzae* herein may be selected from the group of non-typeable strains consisting of strains 11, 33, 32, 29, M4071, K9, K22 and 12. Specific nucleic acid sequences for the genes encoding the respective Hia proteins from such strains are provided herein and are described below.

The nucleic acid molecules provided herein are useful in diagnostic applications. Accordingly, in a further aspect of the invention, there is provided a method of determining the presence, in a sample, of nucleic acid encoding a *Haemophilus influenzae* adhesin protein, comprising the steps of:

a) contacting the sample with a nucleic acid molecule as provided herein to produce duplexes comprising the nucleic acid molecule provided herein are nucleic acid encoding the Hia protein of a strain of *Haemophilus* present in the sample and specifically hybridizable therewith; and

b) determining the production of the duplexes.

In addition, the present invention provides a diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a *Haemophilus influenzae* adhesin protein, comprising:

a) a nucleic acid molecule as provided herein;

b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any such nucleic acid molecule; and

c) means for determining production of the duplexes.

The recombinantly produced truncated Hia proteins provided herein also are useful in diagnostic applications. Accordingly, in another aspect of the invention, there is provided a method of determining the presence of antibodies specifically reactive with the Hia protein in a sample, comprising the steps of (a) contacting the sample with the recombinant Hia protein provided herein to provide complexes of the recombinant Hia protein and any such antibodies present in the sample specifically reactive therewith; and (b) determining production of the complexes.

Advantages of the present invention include:

- an isolated and purified nucleic acid molecule encoding a *Haemophilus influenzae* adhesin protein or a fragment or an analog of the Hia protein;
- recombinantly-produced Hia proteins, free from any other *Haemophilus* proteins; and
- diagnostic kits and immunological reagents for specific identification of *Haemophilus*.

BRIEF DESCRIPTION OF DRAWINGS

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1A shows a restriction map for plasmid DS-2008-2-3 that contains the T7 promoter and the full-length NTHi strain 11 *hia* gene.

Figure 1B shows the oligonucleotides used to PCR amplify the strain 11 *hia* gene. Sense Strand (5038.SL): SEQ ID No: 1, encoded amino acids SEQ ID No: 2; Antisense Strand (5039.SL): SEQ ID No: 3, complement SEQ ID No: 4, encoded amino acids SEQ ID No: 5. Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; H, *Hind* III; N, *Nde* I; Ps, *Pst* I; Sty, *Sty* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance.

Figure 2 shows an immunoblot of the recognition of full-length rHia protein by anti-native *Moraxella catarrhalis* high molecular weight adhesin antibody. Lane 1, DS-2043-1 uninduced; lane 2, DS-2043-1, induced for 4h; lane 3, DS-2043-2 uninduced; lane 4, DS-2043-2, induced for 4h; lane 5, molecular weight markers. DS-2043-1 and DS-2043-2 are independent clones of *pT7 hia(11)* in BL21 (DE3).

Figure 3 shows the construction of plasmids DS-2092-1 and DS-2092-40 that contain tandem copies of the *T7 hia* gene cassette for the strain 11 *hia* gene. Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; H, *Hind* III; Ps, *Pst* I; Xb, *Xba* I. Other abbreviations are: CAP, calf alkaline phosphatase; T7p, T7 promoter; ApR, ampicillin resistance.

Figure 4 shows the sites of truncation for the strain 11 Hia protein (SEQ ID No: 6).

Figure 5A shows the construction of plasmids expressing truncated *hia* genes from strain 11.

Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; H, *Hind* III; N, *Nde* I; Nhe, *Nhe* I; Ps, *Pst* I; R, *EcoR* I; Sty, Sty I; Xb, *Xba* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance.

Figure 5B shows the oligonucleotides used to PCR amplify the 5'-fragments for the truncated genes. E21 truncation: Sense (5524.SL): SEQ ID No: 7, encoded amino acids SEQ ID No: 8; T33 truncation: Sense (5525.SL) SEQ ID No: 9, encoded amino acids SEQ ID No: 10; V38 truncation: Sense (5526.SL): SEQ ID No: 11, encoded amino acids, SEQ ID No: 12; N52 truncation: Sense (5527.SL): SEQ ID No: 13, encoded amino acids SEQ ID No: 14; Antisense (5528.SL): SEQ ID No: 15; complement SEQ ID No: 16, encoded amino acids SEQ ID No: 17.

Figure 6A shows the construction of plasmid BK-96-2-11 that contains the V38 *hia* gene from NTHi strain 11 and the *E. coli cer* gene. Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; K, *Kpn* I; N, *Nde* I; P, *Pst* I; R, *EcoR* I; S, *Sal* I; Sm, *Sma* I; Sty, *Sty* I; Xb, *Xba* I; Xho, *Xho* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance; CAP, calf alkaline phosphatase; tt1 transcription terminator 1 from *trpA*; tt2, transcription terminator 2 from T7 gene 10.

Figure 6B shows the oligonucleotides used to construct the multiple cloning site and transcription terminators. "R" and "Ps" indicate termini that will overlap with EcoR I or Pst I ends, but will not re-

generate the sites. Upperstrand (SEQ ID No.: 50) lower strand (SEQ ID No.: 51).

Figure 7A shows the construction of plasmids DS-2242-1 and DS-2242-2 that contain the T7 promoter and full-length NTHi strain 33 *hia* gene, the *E. coli cer* gene and the kanamycin resistance gene. Restriction enzyme sites are: A, *AlwN* I; B, *BamH* I; Bg, *Bgl* II; H, *Hind* III; K, *Kpn* I; N, *Nde* I; Ps, *Pst* I; R, *EcoR* I; S, *Sal* I; Sm, *Sma* I; Xb, *Xba* I; Xho, *Xho* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance; tt1, transcription terminator 1 from *trpA*; tt2, transcription terminator 2 from T7 gene 10.

Figure 7B shows the oligonucleotides used to generate the 5'-end of the strain 33 *hia* gene coding strand (SEQ ID No.: 52), complementary strand (SEQ ID No.: 53), and encoded amino acid sequence (SEQ ID No.: 54).

Figure 8A shows the construction of plasmid DS-2340-2-3 that contains the T7 promoter and the V38 *hia* gene from strain 33, the E. coli *cer* gene and the kanamycin resistance gene. Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; H, *Hind* III; N, *Nde* I; Ps; *Pst* I; R, *EcoR* I; S, *Sal* I; Sn, *SnaB* I; Xb, *Xba* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance; tt1, transcription terminator 1 from *trpA*; tt2, transcription terminator 2 from T7 gene 10.

Figure 8B shows the oligonucleotides used to PCR amplify the 5'-end of the truncated *hia* gene. Sense (6286.8L): SEQ ID No: 60, encoded amino acids SEQ ID

No: 61; antisense (6287.SL) SEQ ID No: 18, complement SEQ ID No: 19, encoded amino acids SEQ ID No: 20.

Figures 9A and 9B show the construction of plasmids DS-2447-2 and DS-2448-17, that contain tandem copies of the *T7 V38 hia (11)* and *T7 V38 hia (33)* genes, respectively. Restriction enzyme sites are: B, BamH I; Bg, Bgl II; H, Hind III; Ps, Pst I; R, EcoR I; S, Sal I; Xb, Xba I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance; CAP, calf alkaline phosphatase; tt1, transcription terminator 1 from trpA; tt2, transcription terminator 2 from T7 gene 10.

Figure 10 shows the expression of rHia. Panel A: lane 1, full-length rHia (11) no induction; lane 2, full-length rHia (11); lane 3, E21 rHia (11); lane 4, T33 rHia (11); lane 5, V38 rHia (11); lane 6, N52 rHia (11). Panel B: lane 1, V38 rHia (11) no induction; lane 2, V38 rHia (11); lane 3, V38 rHia (11)/cer.

Figure 11 shows a purification scheme for rHia proteins. Abbreviations are: SP, supernatant; PPT, precipitate; DTT, dithiothreitol; OG, octyl glucoside; (x) means discarded.

Figure 12, having panels A and B, shows the SDS-PAGE analysis of purified rHia. Panel A shows purified V38 rHia protein from strain 11 and panel B shows purified V38 rHia protein from strain 33. Lane 1, molecular weight markers; lane 2, whole-cell lysate; lane 3, crude extract; lane 4, purified rHia protein.

Figure 13, having panels A, B and C, shows the stability of V38 rHia (11). Panel A shows samples stored at 4°C without glycerol. Panel B shows samples

stored at 4°C, in the presence of 20% glycerol. Panel C shows samples stored at -20°C in the presence of 20% glycerol. Lane 0 indicates t_0; lanes 1 to 8 indicate samples stored for 1 to 8 weeks.

Figure 14, having panels A and B, shows the immunogenicity of V38 rHia (11) or V38 rHia (33) in CD-1 mice. Panel A shows the response after a single immunization and panel B shows the response of a prime/boost immunization.

Figures 15A and 15B show the immunogenicity of V38 rHia (11) in BALB/c mice and guinea pigs. Figure 15A shows the antibody response in mice and Figure 15B shows the response in guinea pigs.

Figure 16 illustrates the protective ability of V38 rHia (33) against nasopharyngeal colonization in a chinchilla model.

Figure 17 shows the oligonucleotides used to PCR amplify additional *hia* genes. Sense (5040.SL), SEQ ID No: 21, encoded amino acids SEQ ID No: 22; Antisense (5039.SL), SEQ ID No: 3, complement SEQ ID No: 4, encoded amino acids SEQ ID No: 5.

Figure 18 shows the nucleotide sequence (SEQ ID No: 23) and deduced amino acid sequence (SEQ ID No: 24) of the *hia* gene from NTHi strain 33.

Figure 19 shows the nucleotide sequence (SEQ ID No: 25) and deduced amino acid sequence (SEQ ID No: 26) of the hia gene from NTHi strain 32.

Figure 20 shows the nucleotide sequence (SEQ ID No: 27) and deduced amino acid sequence (SEQ ID No: 28) of the *hia* gene from NTHi strain 29.

Figure 21 shows the nucleotide sequence (SEQ ID No: 29) and deduced amino acid sequence (SEQ ID No: 30 of the *hia* gene from NTHi strain M4071.

Figure 22 shows the nucleotide sequence (SEQ ID No: 31) and deduced amino acid sequence (SEQ ID No: 32) of the *hia* gene from NTHi strain K9.

Figure 23 shows the nucleotide sequence (SEQ ID No: 33) and deduced amino acid sequence (SEQ ID No: 34) of the *hia* gene from NTHi strain K22.

Figure 24 shows the nucleotide sequence (SEQ ID No: 35) and deduced amino acid sequence (SEQ ID No: 36) of the *hia* gene from type c strain API.

Figure 25 shows the nucleotide sequence (SEQ ID No: 37) and deduced amino acid sequence (SEQ ID No: 38) of the *hia* locus from NTHi strain 12. The overlined or underlined sequences indicate oligonucleotides used to PCR amplify across the junction of the two *orfs*. Sense (6431.SL) SEQ ID No: 39, (6432.SL) SEQ ID No: 40; antisense (6295.SL) SEQ ID No: 41, (6271.SL) SEQ ID No: 42.

Figure 26 shows the nucleotide sequence (SEQ ID No.: 43) and deduced amino acid sequence (SEQ ID No.: 44) of the *hia* locus from NTHi strain 11, as published in U.S. Patent No. 5,646,259.

Figure 27 shows the alignment of the upstream ORF from the strain 12 *hia* locus (SEQ ID No: 45) with part of the HI1732 protein (SEQ ID No: 46) from *H. influenzae* type b strain Rd.

Figure 28 shows the alignment of amino acid sequences from Hia (SEQ ID Nos. 24, 26, 28, 34, 30, 44, 32), Hsf (SEQ ID No.: 47) and partial sequences from

Moraxella catarrhalis high molecular weight proteins (200 kDa) from strains 4223 and LES-1 (SEQ ID Nos.: 48, 49). Asterisks within sequences indicate stop codons, but below the sequence they indicated sequence homology. Dots indicate identical residues. The sequence alignments were prepared by direct comparison of the amino acid sequences of the respective proteins.

Figure 29 shows the oligonucleotides used to PCR amplify the 5' end of the *hia* gene at the S44 truncated position. Sense (6817.SL) SEQ ID No: 55, encoding amino acids SEQ ID No: 56; antisense (6818.SL) SEQ ID No: 57, complement SEQ ID No: 58, encoded amino acids SEQ ID No: 59.

Figure 30 shows the construction of plasmid JB-2930-3 that contains the S44 *hia* gene from NTHi strain 11 and the E. coli *cat* gene and the T7 promoter. Restriction enzyme sites are: B, *BamH* I; Bg, *Bgl* II; K, *Kpn* I; N, *Nde* I; P, *Pst* I; R, *EcoR* I; S, *Sal* I; Sm, *Sma* I; Sty, *Sty* I; Xb, *Xba* I; Xho, *Xho* I. Other abbreviations are: T7p, T7 promoter; ApR, ampicillin resistance; KanR, kanamycin resistance; CAP, calf alkaline phosphatase; tt1 transcription terminator 1 from *trpA*; tt2, transcription terminator 2 from T7 gene 10.

Figure 31 shows SDS-PAGE analysis of the expression of rHia from S44. Lane 1, expression from pET S44 vector at time 0 (no induction); lane 2 expression from pET S44 vector after 4 hours induction; lane 3 expression from JB-2930-3 after 4 hours induction.

Figures 32 shows a schematic representation of the two vectors used for the expression study, JB-2930-3 and IA-191-3-1, of S44-truncated rHia.

GENERAL DESCRIPTION OF THE INVENTION

Since *H. influenzae* strains produce low quantities of the Hia and Hsf proteins, the *hia* gene from NTHi strains was cloned into an expression vector for overproduction of the recombinant protein in *E. coli*. When the full-length recombinant Hia (rHia) protein was expressed, it was made in relatively low quantities. In order to confirm that there was expression of the recombinant protein, an immunoblot was performed using antibody raised to a *Moraxella catarrhalis* high molecular weight adhesin protein identified as 200 kDa in US Patent No. 5,808,024, assigned to the assignee and the disclosure of which is incorporated herein by reference. Antibody against the gel-purified native 200 kDa protein recognized a specific induced band in the rHia protein sample. The yield of rHia was not significantly improved by increasing the gene copy number of the T7 *hia* gene cassette.

The *E. coli cer* gene has been shown to stabilize plasmids containing large inserts (ref. 15), but the yield of rHia was not significantly improved by adding the *E. coli cer* gene to the expression vector. However, the *E. coli* cells were observed to clump during culture, suggesting that there was surface expression of the Hia adhesin protein. The apparent toxicity of the rHia protein might be overcome if it were made as inclusion bodies, so truncations were made at the 5'-end of the gene to delete putative signal

sequences. This modification resulted in good production and recovery of truncated rHia starting from the V38 position.

The full-length and V38-truncated rHia proteins were immunogenic and the resultant anti-rHia antibodies were protective in passive infant rat models of bacteremia due to *H. influenzae* type a or type b strains. In addition, the truncated V38 rHia protein was found to be partially protective against nasopharyngeal colonization in an active challenge model in chinchillas. The protection afforded by rHia derived from an NTHi strain against disease caused by NTHi and encapsulated type a or type b strains, indicates that there may be common protective epitopes. The cloning and sequence analysis of additional *hia* genes may help to identify conserved regions. The full-length or N-terminal truncated rHia proteins may be used as vaccine components to protect against *Haemophilus influenzae* disease.

Any *Haemophilus* strains that have *hia* genes may be conveniently used to provide the purified and isolated nucleic acid molecules (which may be in the form of DNA molecules), comprising at least a portion coding for a Hia protein as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as American Type Culture Collection. Appropriate strains of *Haemophilus* include:

Non-typeable *Haemophilus* strain 11;

Non-typeable *Haemophilus* strain 33;

Non-typeable *Haemophilus* strain 32;

Non-typeable *Haemophilus* strain 29;

Non-typeable *Haemophilus* strain M4071;

Non-typeable *Haemophilus* strain K9;

Non-typeable *Haemophilus* strain K22;

Non-typeable *Haemophilus* strain 12;

Type C *Haemophilus* strain API.

In this application, the term "Hia" protein is used to define a family of Hia proteins that includes those having naturally occurring variations in their amino acid sequences as found in various strains of *Haemophilus*.

Referring to Fig. 1A, there is illustrated a restriction map of plasmid DS-2008-2-3 that contains a full-length *hia* gene from non-typeable *Haemophilus influenzae* strain 11, under the influence of the T7 promoter. The nucleic acid (SEQ ID No.: 43) and deduced amino acid sequence (SEQ ID No.: 44) of the *hia* gene from strain 11, are described in the aforementioned U.S. Patent No. 5,646,259 (and identified therein as "HA1"). The oligonucleotides used to PCR amplify the *hia* gene from the ATG start codon of the gene of strain 11 are shown in Fig. 1B.

Referring to Fig. 2, there is illustrated an immunoblot demonstrating the recognition of the rHia (11) protein by anti-native *Moraxella catarrhalis* high molecular weight adhesin antibody. The *M. catarrhalis* high molecular weight adhesin or 200 kDa protein described in the aforementioned US Patent No. 5,808,024 has some sequence homology with the Hia and Hsf proteins, especially at the carboxy terminus (Fig. 28).

Referring to Fig. 3, there is illustrated a construction scheme for plasmids DS-2092-1 and DS-2092-40 that contain tandem copies of T7 *hia* gene cassettes comprising the full-length *hia* gene from NTHi strain 11. Such plasmids that contain increased copy numbers of genes often have enhanced production levels for recombinant proteins. However, as seen below, the low yield of recombinant Hia was not significantly improved by increasing the gene copy number.

Referring to Fig. 4, there is illustrated the N-terminal sequence of the NTHi strain 11 protein and the position of time N-terminally truncated rHia proteins. The N-terminal truncation up to position E21 deletes a long hydrophobic region that may constitute part of a signal sequence for Hia. The deletion up to position T33 includes a long hydrophobic region and follows a potential Ala-X-Ala signal cleavage site. The deletion up to position V38 includes a long hydrophobic region and follows a potential Ala-X-Ala signal cleavage site. The recombinant Hia protein starting at position S44 includes a long hydrophobic region and follows a potential Ala-X-Ala signal cleavage site. The recombinant Hia protein starting at position N52 mimics the approximate start of the related high molecular weight (200 kDa) adhesin from *Moraxella catarrhalis* described in the aforementioned US Patent No. 5,808,024, which recombinant protein is over-produced if truncated at its N-terminus to start at V56.

Referring to Fig. 5A, there is illustrated the construction scheme for the generation of plasmids DS-2186-1-1, DS-2201-1, DS-2186-2-1, and DS-2168-2-6

producing four of the N-terminal truncated rHia proteins. The oligonucleotides used to PCR amplify the 5'-fragments are shown in Fig. 5B. In Figure 30, there is illustrated the construction scheme for the generation of plasmids JB-2930-3, which produces the S44 deletion. The oligonucleotides used to PCR amplify the 5'-fragments are shown in Figure 29.

Referring to Fig. 6A, there is illustrated a construction scheme for the generation of plasmid BK-96-2-11 that contains the V38 *hia* gene from NTHi strain 11 as well as the *E. coli cer* gene that has been shown to stabilize plasmids. The introduction of the *cer* gene into plasmids producing toxic proteins, was predicted to enhance protein production. There was an observed change in the morphology of the *E. coli* cells producing full-length rHia in the presence of the *cer* gene, in that they clumped. This suggests that there was enhanced expression of the adhesin at the surface of the cells that caused the clumping. The expression plasmid BK-96-2-11 also contains transcription terminators upstream and downstream of the T7 V38 *hia* gene cassette that were predicted to enhance the gene stability. The oligonucleotides used to generate the multiple cloning site and transcription terminators are shown in Fig. 6B.

Referring to Fig. 7A, there is illustrated a construction scheme for plasmids DS-2242-1 and DS-2242-2 that contain a full-length *hia* gene from non-typeable *Haemophilus influenzae* strain 33, under the influence of the T7 promoter. The expression plasmids also contain the *E. coli cer* gene and transcription

terminators upstream and downstream of the T7 *hia* (33) gene cassette. DS-2242-1 has the terminators coded on the same strand as the T7 *hia* (33) gene. However, there was no observable difference in the expression of rHia from the two plasmids. The oligonucleotides used to construct the authentic 5'-end of the NTHi strain 33 gene are shown in Fig. 7B.

Referring to Fig. 8A, there is illustrated a construction scheme for plasmid DS-2340-2-3 that contains the V38 *hia* gene from NTHi strain 33 as well as the *E. coli cer* gene. There are also transcription terminators located upstream and downstream of the T7 V38 *hia* gene cassette, on the same strand. The oligonucleotides used to PCR amplify the NTHi strain 33 *hia* gene from the V38 codon, are shown in Fig. 8B.

Referring to Fig. 9, there is shown the construction of plasmids DS-2447-2 and DS-2448-17 that contain tandem copies of the *T7 V38 hia (11)* or *T7 V38 hia (33)* gene cassettes, respectively.

Referring to Fig. 10, panel A, there is illustrated the production of rHia proteins from plasmids encoding full-length or truncated *hia* genes from NTHi strain 11. The production of the full-length rHia (11) protein was very low. There was also low expression observed for the E21 and T33 truncated rHia proteins. However, the V38 and N52 truncated rHia proteins have significantly improved expression levels. As shown in Fig. 10, panel B, the production of V38 rHia (11) appears to be enhanced when the *E. coli cer* gene is added to the expression plasmid.

Referring to Fig. 11, there is illustrated a purification scheme for rHia proteins, produced as inclusion bodies. Cells were lysed by sonication and the inclusion bodies purified by serial extractions. The inclusion bodies were solubilized in guanidinium chloride and impurities precipitated by the addition of polyethlyene glycol (PEG). Addition of $(NH_4)_2SO_4$ resulted in precipitation of rHia and the crude rHia was further purified by gel filtration.

Referring to Fig. 12, there is illustrated the purified V38 rHia proteins from strains 11 and 33. The inclusion bodies are shown in lane 3 and the final purified protein in lane 4. The estimated purity of the purified protein is greater than about 90% as determined by SDS-PAGE densitometry.

Referring to Fig. 13, there is shown the SDS-PAGE analysis of the stability of rHia proteins produced as described herein during 8 weeks of storage with or without glycerol at 4°C and with glycerol at -20°C. The protein is stable under any of these conditions.

Referring to Fig. 14, there is illustrated the immunogenicity of V38 rHia proteins from strains 11 and 33 in CD-1 mice. At doses from 0.3 to 10 µg, there is a strong immune response after one or two doses with either protein. There is no obvious dose response at these levels. Similar results were observed in BALB/c mice (Fig. 15A) and in guinea pigs (Fig. 15B), indicating that rHia was very immunogenic, even at 0.3 µg per dose.

Referring to Fig. 16, there is illustrated the protection afforded by V38 rHia (33) against

colonization by NTHi strain 33. As described by Yang et al (ref. 20), a chinchilla nasopharyngeal colonization model has been developed to assess protection against this earliest stage of disease. The model was initially established for NTHi strains that express *hmw* genes and had to be adapted for NTHi strains expressing *hia* genes. For the prototype *hmw*-expressing strain (NTHi 12), 10^2 to 10^8 cfu could be used to establish infection, but 5 x 10^8 cfu of NTHi strain 33 was required, and even at this high level no infection could be established with the prototype *hia*-expressing strain 11. At a 100 µg dose, it is evident that there is partial protection in the immunized cohort, although there is no protection at a 50 µg dose. Such protection against the early stages of disease illustrates the utility of the rHia adhesins as vaccine antigens.

Referring to Fig. 17, there is illustrated the oligonucleotides used to PCR amplify additional *Haemophilus influenzae hia* genes. The sequences are based upon the conserved amino and carboxy terminal sequences of the Hia and Hsf proteins.

Referring to Fig. 18, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the NTHi strain 33 *hia* gene. Referring to Fig. 19, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the NTHi strain 32 *hia* gene. Referring to Fig. 20, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the NTHi strain 29 *hia* gene. Referring to Fig. 21, there is illustrated the

complete nucleotide sequence and deduced amino acid sequence of the NTHi strain M4071 *hia* gene. Referring to Fig. 22, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the NTHi strain K9 *hia* gene. Referring to Fig. 23, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the NTHi strain K22 *hia* gene. Referring to Fig. 24, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the *Haemophilus influenzae* type c strain API *hia* gene. Referring to Fig. 25, there is illustrated the complete nucleotide sequence and deduced amino acid sequence of the *hia* locus from NTHi strain 12. The PCR amplified fragment contains the 3'-end of a gene related to HI1733 gene of the *Haemophilus influenzae* type d strain Rd genome joined to the 3'-end of an *hia* gene. An alignment of the upstream ORF with the HI1733 protein is shown in Fig. 27.

Figure 26 shows the complete nucleotide sequence and the deduced amino acid sequence of the *Hia* gene from NTHi strain 11, as published in the aforementioned USP 5,646,259.

Referring to Fig. 28, there is illustrated an alignment of the deduced protein sequences from Hsf, Hia, and partial sequences of the *M. catarrhalis* 200 kDa protein.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have use in applications in the fields of vaccination, diagnosis, treatment of *Haemophilus* infection and the generation of immunological agents. A further non-

limiting discussion of such uses is further presented below.

Vaccine Preparation and Use

Immunogenic compositions, suitable to be used as vaccines, may be prepared from immunogenic recombinant *Haemophilus influenzae* adhesin (rHia) proteins of non-typeable *Haemophilus* strains, immunogenic analogs and fragments thereof and/or immunogenic peptides as disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-rHia antibodies and antibodies that are opsonizing or bactericidal.

Immunogenic compositions, including vaccines, may be prepared as injectables, as liquid solutions or emulsions. The rHia protein, immunogenic analogs and fragments thereof and/or immunogenic peptides may be mixed with pharmaceutically acceptable excipients which are compatible with the rHia protein, immunogenic fragments analogs or immunogenic peptides. Such excipients may include, water, saline, dextrose, glycerol, ethanol and combinations thereof.

The immunogenic compositions and vaccines may further contain auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants to enhance the effectiveness of the vaccines.

Immunogenic compositions and vaccines may be administered parenterally, by injection subcutaneously or intramuscularly. Alternatively, the immunogenic compositions formed according to the present invention, may be formulated and delivered in a manner to evoke an immune response at mucosal surfaces. Thus, the

immunogenic composition may be administered to mucosal surfaces by, for example, the nasal or oral (intragastric) routes.

The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. Some such targeting molecules include vitamin B12 and fragments of bacterial toxins, as described in WO 92/17167 (Biotech Australia Pty. Ltd.), and monoclonal antibodies, as described in U.S. Patent No. 5,194,254 (Barber et al).

Alternatively, other modes of administration including suppositories and oral formulations may be desirable. For suppositories, binders and carriers may include, for example polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example pharmaceutical grades of saccharine, cellulose and magnesium carbonate. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 1 to 95% of the rHia protein, fragment analogs and/or peptides.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective, protective and immunogenic. The quantity to be administered depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and if needed, to produce a cell-mediated immune response. Precise amounts of

active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the rHia, analogs and fragments thereof and/or peptides. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and will vary according to the size of the host.

The nucleic acid molecules encoding the rHia proteins of non-typeable *Haemophilus* may also be used directly for immunization by administration of the DNA directly, for example by injection for genetic immunization or by constructing a live vector, such as *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus, containing the nucleic acid molecule. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system is contained in, for example, O'Hagan (1992) (ref. 16). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulmer et al., 1993 (ref. 17).

Immunogenicity can be significantly improved if the antigens are co-administered with adjuvants, commonly used as an 0.05 to 1.0 percent solution in phosphate - buffered saline. Adjuvants enhance the immunogenicity of an antigen but are not necessarily immunogenic themselves. Adjuvants may act by retaining the antigen locally near the site of administration to

produce a depot effect facilitating a slow, sustained release of antigen to cells of the immune system. Adjuvants can also attract cells of the immune system to an antigen depot and stimulate such cells to elicit immune responses.

Immunostimulatory agents or adjuvants have been used for many years to improve the host immune responses to, for example, vaccines. Intrinsic adjuvants, such as lipopolysaccharides, normally are the components of the killed or attenuated bacteria used as vaccines. Extrinsic adjuvants are immunomodulators which are typically non-covalently linked to antigens and are formulated to enhance the host immune responses. Thus, adjuvants have been identified that enhance the immune response to antigens delivered parenterally. Some of these adjuvants are toxic, however, and can cause undesirable side-effects, making them unsuitable for use in humans and many animals. Indeed, only aluminum hydroxide and aluminum phosphate (collectively commonly referred to as alum) are routinely used as adjuvants in human and veterinary vaccines. The efficacy of alum in increasing antibody responses to diphtheria and tetanus toxoids is well established.

A wide range of extrinsic adjuvants can provoke potent immune responses to antigens. These include the specific adjuvants detailed above as well as saponins complexed to membrane protein antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria and mineral oil, Freund's complete adjuvants, bacterial products, such as muramyl

dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, and liposomes.

To efficiently induce humoral immune responses (HIR) and cell-mediated immunity (CMI), immunogens are emulsified in adjuvants. Many adjuvants are toxic, inducing granulomas, acute and chronic inflammations (Freund's complete adjuvant, FCA), cytolysis (saponins and pluronic polymers) and pyrogenicity, arthritis and anterior uveitis (LPS and MDP). Although FCA is an excellent adjuvant and widely used in research, it is not licensed for use in human or veterinary vaccines because of its toxicity.

Desirable characteristics of ideal adjuvants include:

(1) lack of toxicity;

(2) ability to stimulate a long-lasting immune response;

(3) simplicity of manufacture and stability in long-term storage;

(4) ability to elicit both CMI and HIR to antigens administered by various routes, if required;

(5) synergy with other adjuvants;

(6) capability of selectively interacting with populations of antigen presenting cells (APC);

(7) ability to specifically elicit appropriate T_H1 or T_H2 cell-specific immune responses; and

(8) ability to selectively increase appropriate antibody isotype levels (for example, IgA) against antigens.

US Patent No. 4,855,283 granted to Lockhoff et al on August 8, 1989 which is incorporated herein by

reference thereto teaches glycolipid analogues including N-glycosylamides, N-glycosylureas and N-glycosylcarbamates, each of which is substituted in the sugar residue by an amino acid, as immuno-modulators or adjuvants. Thus, Lockhoff et al. 1991 (ref. 18) reported that N-glycolipid analogs displaying structural similarities to the naturally-occurring glycolipids, such as glycosphingolipids and glycoglycerolipids, are capable of eliciting strong immune responses in both herpes simplex virus vaccine and pseudorabies virus vaccine. Some glycolipids have been synthesized from long chain-alkylamines and fatty acids that are linked directly with the sugars through the anomeric carbon atom, to mimic the functions of the naturally occurring lipid residues.

U.S. Patent No. 4,258,029 granted to Moloney, assigned to the assignee hereof and incorporated herein by reference thereto, teaches that octadecyl tyrosine hydrochloride (OTH) functions as an adjuvant when complexed with tetanus toxoid and formalin inactivated type I, II and III poliomyelitis virus vaccine. Also, Nixon-George et al. 1990 (ref. 19), reported that octadecyl esters of aromatic amino acids complexed with a recombinant hepatitis B surface antigen, enhanced the host immune responses against hepatitis B virus.

Immunoassays

The rHia protein of a non-typeable strain of *Haemophilus*, analogs and fragments thereof produced according to the present invention are useful as immunogens, as antigens in immunoassays including enzyme-linked immunosorbent assay (ELISA), RIAs and

other non-enzyme linked antibody binding assays or procedures known in the art for the detection of anti-bacterial, *Haemophilus*, and/or Hia antibodies. In ELISA assays, the Hia protein, analogs and fragments are immobilized onto a selected surface, for example a surface capable of binding proteins or peptides, such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed Hia protein, analogs and/or fragments, a nonspecific protein such as a solution of bovine serum albumin (BSA) or casein that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by nonspecific bindings of antisera onto the surface.

The immobilizing surface is then contacted with a sample, such as clinical or biological materials, to be tested in a manner conducive to immune complex (antigen/antibody) formation. This may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to incubate for from about 2 to about 4 hours, at temperature such as of the order of about 25° to about 37°C. Following incubation, the sample-contacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween, or a borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound Hia protein,

analogs and/or fragments, and subsequent washing, the occurrence, and even amount, of immunocomplex formation may be determined by subjecting the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and in general IgG. To provide detecting means, the second antibody may have an associated activity, such as an enzymatic activity, that will generate, for example, a color development, upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for example, a visible spectra spectrophotometer.

Use of Sequences as Hybridization Probes

The nucleotide sequences of the present invention, comprising the newly-isolated and characterized sequences of the *hia* genes, allow for the identification and cloning of the *hia* genes from other non-typeable strains of *Haemophilus*.

The nucleotide sequences comprising the sequence of *hia* genes of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other *hia* genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other *hia* genes in other strains of non-typeable *Haemophilus*. For a high degree of selectivity, relatively stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as

provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amount of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results. In general, convenient hybridization temperatures in the presence of 50% formamide and 0.15 M NaCl are: 42°C for an *hia* gene which is about 95 to 100% homologous to the target nucleic acid fragment, 37°C for about 90 to 95 homology and 32°C for about 85 to 90% homology.

In a clinical diagnostic embodiment, the nucleic acid sequences of the *hia* genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific

hybridization with samples containing *Hia* genes sequences.

The nucleic acid sequences of *Hia* genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid-phase procedures the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e.g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the *hia* genes or fragments thereof of the present invention under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required depending on, for example, the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label. It is preferred to select nucleic acid sequence portions which are conserved among species of *Haemophilus*. The selected probe may be at least 18 bp in length and may be in the range of 30 bp to 90 bp long.

Expression of the *Haemophilus influenzae* adhesin Genes

Plasmid vectors containing replicon and control sequences which are derived from species compatible

with the host cell may be used for the expression of the *hia* genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the host cell for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM™-11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as *E. coli* LE392.

Promoters commonly used in recombinant DNA construction include the β-lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system employed herein in preferred embodiments (U.S. Patent 4,952,496). Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with genes. The particular promoter used will generally be a matter of choice depending upon the desired results. Hosts that are appropriate for expression of the Hia protein and immunological fragments or analogs thereof include *E.*

coli, *Bordetella* species, *Bacillus* species, *Haemophilus*, fungi, yeast or the baculovirus expression system may be used. *E. coli* is the preferred host used herein.

In accordance with this invention, it is preferred to produce the Hia proteins by recombinant methods, particularly when the naturally occurring Hia protein as purified from a culture of a species of *Haemophilus* may include trace amounts of toxic materials or other contaminants. This problem can be avoided by using recombinantly produced Hia protein in heterologous systems which can be isolated from the host in a manner to minimize contaminants in the purified materials, specifically employing the constructs described herein.

BIOLOGICAL DEPOSITS

A vector that contains nucleic acid coding for a high molecular weight protein of a non-typeable strain of *Haemophilus* that is described and referred to herein has been deposited with the America Type Culture Collection (ATCC) located at 10801 University Boulevard, Manassas, Virginia 20110-2209, USA, pursuant the Budapest Treaty and prior to the filing of this application. Samples of the deposited vector will become available to the public and all restrictions imposed or access to the deposits will be received upon grant of a patent based on this United States patent application. In addition, the deposit will be replaced if viable samples cannot be dispensed by the Depository. The invention described and claimed herein is not limited in scope by the biological materials deposited, since the deposited embodiment is intended

only as an illustration of the invention. Any equivalent or similar vectors that contain nucleic acid which encodes equivalent or similar antigens as described in this application are within the scope of the invention.

Deposit Summary

Plasmid	ATCC	Deposit Date
BK-96-2-11	203771	February 11, 1999

EXAMPLES

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific Examples. These Examples are described solely for purposes of illustration and are not intended to limit the scope of the invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations.

Methods of molecular genetics, protein biochemistry, immunology and fermentation technology used, but not explicitly described in this disclosure and these Examples, are amply reported in the scientific literature and are well within the ability of those skilled in the art.

Example 1

This Example describes the construction of plasmid DS-2008-2-3 that expresses full-length rHia proteins from NTHi strain 11.

Chromosomal DNA was purified from NTHi strain 11 and the full-length *hia* gene was PCR amplified using the oligonucleotides (5038.SL and 5039.SL) described in Figure 1B. An Nde I site was engineered at the 5'-end of the gene and a BamH I site was engineered at the 3'-end for cloning into the pT7-7 expression vector (ref. 21). The amplified fragment was digested with Nde I/BamH I and cloned into pT7-7 that had been digested with the same enzymes. Plasmid DS-2008-2-3 contains a 3.4 kb strain 11 *hia* gene downstream of the T7 promoter (Fig. 1A). The plasmid was used to express recombinant Hia (Example 9 below).

Example 2

This Example illustrates the recognition of rHia by anti-native *Moraxella catarrhalis* high molecular weight adhesin antibody.

There is some sequence conservation observed between the *Haemophilus influenzae* Hia proteins and a *Moraxella catarrhalis* high molecular weight adhesin identified as the *M. catarrhalis* 200 kDa protein in aforementioned US Patent No. 5,808,024 (Fig. 28). The native *M. catarrhalis* 200 kDa protein was gel purified as described in US Patent No. 5,808,024 and guinea pig anti-native 200 kDa antibody was generated. The T7 *hia* gene was expressed from plasmid DS-2008-2-3 and the cell culture containing the rHia protein was electroblotted to nitrocellulose membrane. Immunoblot analysis using anti-native 200 kDa antibody showed that the antibody recognized the rHia protein, as seen in Figure 2.

Example 3

This Example describes the construction of plasmids DS-2092-1 and DS-2092-40 that contain tandem copies of *T7 hia* (11) gene cassettes.

In order to improve the production of full-length recombinant Hia protein, tandem copies of the *T7 hia* gene cassette containing the strain 11 *hia* gene (Example 1) were inserted into a single vector. Plasmid DS-2008-2-3 was linearized with *Bgl* II and dephosphorylated. Plasmid DS-2008-2-3 was also digested with *Bgl* II and *BamH* I to excise the *T7 hia* gene cassette. The *T7 hia* fragment was ligated into the linearized vector to generate plasmid DS-2092-1 that contains two copies of the *T7 hia* gene in the anti-clockwise orientation (a,a) and plasmid DS-2092-40 that contains tandem copies in opposite orientations (a,c) (Fig. 3). There was no obvious improvement in expression of rHia from either construct (see Example 9 below).

Example 4

This Example describes the construction of plasmids expressing truncated strain 11 *hia* genes.

The production of the rHia protein from single or tandem copies of the *T7 hia* gene cassette was very low and the protein seemed to be toxic to *E. coli* (as described below in Example 9). Since *H. influenzae* Hia is a surface-exposed adhesin molecule, it must either utilize a signal sequence or accessory protein(s) for secretion, but there are no known accessory genes involved. If the signal sequence were removed for expression of the recombinant protein in *E. coli*, the

rHia might be expressed as inclusion bodies and the toxic effect reduced. A putative signal sequence and cleavage sites were identified and four constructs expressing N-terminally truncated rHia proteins were designed (Fig. 4). There is a unique *Sty* I site in the strain 11 *hia* gene about 500 bp from the start codon. Plasmid DS-2008-2-3 was digested with *Nde* I and *Sty* I and the 5.7 kb vector fragment purified (Fig. 5A). PCR primers were designed to amplify from the truncation site to the *Sty* I site and a unique *Nhe* I site was introduced into the antisense primer for screening truncated clones (Fig. 5B). The amplified fragments were subcloned into pCRII for easier manipulation, generating plasmids DS-2153R-1-2 (E21), DS-2165-4-8 (T33), DS-2153-3-5 (V38), and DS-2153-4-4 (N52). The pCRII *hia* plasmids were digested with *Nde* I and *Sty* I and the fragments ligated with the vector piece from DS-2008-2-3. Plasmids DS-2186-1-1 (E21), DS-2201-1 (T33), DS-2186-2-1 (V38), and DS-2168-2-6 (N52) were generated that contained the T7 promoter and truncated *hia* genes as indicated in parentheses. These plasmids were used to express recombinant Hia (see Example 9 below).

Example 5

This Example describes the construction of plasmid BK-96-2-11 that contains the *T7 V38 hia (11)* cassette, the *E. coli cer* gene, and the kanamycin resistance gene.

Plasmid DS-1843-2 is a pBR328-based plasmid in which a multiple cloning site and two transcription terminators have been introduced on oligonucleotides,

between the *EcoR* I and *Pst* I sites, thus destroying both the chloramphenicol and ampicillin resistance genes (Fig. 6B). The kanamycin resistance gene from pUC-4K was inserted at the *Sal* I site, to generate plasmid DS-2147-1 that is kanamycin resistant and tetracycline sensitive. Plasmid DS-2224-1-4 is a pUC plasmid containing a synthetic *E. coli cer* gene (ref. 15) constructed from oligonucleotides and flanked by *BamH* I sites. The 290 bp *BamH* I fragment of the *cer* gene was inserted into the *BamH* I site of DS-2147-1 creating plasmid BK-2-1-2. This pBR-based plasmid thus contains a multiple cloning site, the kanamycin resistance gene and the *cer* gene. Plasmid BK-2-1-2 was linearized with *Bgl* II and dephosphorylated. Plasmid DS-2186-2-1 was digested with *Bgl* II and *BamH* I and the 3.6 kb *T7 V38 hia* fragment was inserted into BK-2-1-2, creating plasmid BK-96-2-11 (Fig. 6A).

Example 6

This Example describes the construction of plasmids DS-2242-1 and DS-2242-2 that express the full-length NTHi strain 33 *hia* gene in the presence of the *E. coli cer* gene.

Chromosomal DNA was purified from NTHi strain 33 and PCR amplification was performed using oligonucleotides 5039.SL and 5040.SL (Fig. 17). The sense primer (5040.SL) was designed based upon the 5'-flanking sequence of strain 11 *hia* and the conserved amino terminal sequences of the NTHi Hia and Hib Hsf proteins. The antisense primer (5039.SL) was the same as that described in Example 1 and was based upon the conserved carboxy terminal sequences of the Hia and Hsf

proteins. The 3 kb strain 33 *hia* PCR fragment was cloned into pCR II, generating plasmid DS-1917-3-8.

In order to express the full-length strain 33 *hia* gene, approximately 106 bp of the 5'-end of the gene was synthesized from oligonucleotides, from the start codon to an *AlwN* I site (Fig. 7B). Plasmid DS-1917-3-8 was digested with *AlwN* I and *BamH* I and the approximately 2.9 kb fragment containing the *hia* gene was purified. Plasmid pT7-7 was digested with *Nde* I and *BamH* I. The *Nde* I - *AlwN* I oligonucleotides and *AlwN* I - *BamH* I *hia* fragment were ligated into the pT7-7 vector, generating plasmid DS-2103-4.

In order to include the *E. coli cer* gene and utilize kanamycin selection, the *Bgl* II - *BamH* I fragment containing the *T7 hia (33)* gene cassette was excised from DS-2103-4 and cloned into BK-2-1-1 that had been digested with *Bgl* II and dephosphorylated. Plasmids DS-2242-1 and DS-2242-2 contain single copies of the *T7 hia (33)* gene cassette in opposite orientations, the *E. coli cer* gene, and the kanamycin resistance gene (Fig. 7A).

Example 7

This Example describes the construction of plasmid DS-2340-2-3 that contains a *T7 hia* gene cassette with a truncated V38 strain 33 *hia* gene, the *E. coli cer* gene, and the kanamycin resistance gene.

PCR primers were designed to amplify a 250 bp fragment of the 5'-end of the NTHi strain 33 hia gene from a V38 start codon up to an internal *SnaB* I site. An *Nde* I site was added at the 5'-end for cloning purposes and the fragment was amplified using plasmid

DS-2242-1 as template. The construction scheme is shown in Figure 8A and the PCR primers are shown in Figure 8B. The fragment was cloned into pCR II generating plasmid DS-2328-1-1. DS-2242-1 was digested with *Nde* I and *SnaB* I and the 8.5 kb vector fragment purified. DS-2328-1-1 was digested with *Nde* I and *SnaB* I and the 0.25 kb *5' hia* fragment was ligated with the 8.5 kb vector fragment from DS-2242-1, to generate plasmid DS-2340-2-3.

Example 8

This Example illustrates the construction of plasmids DS-2447-2 and DS-2448-17 that contain tandem copies of *T7 V38 hia (11)* or *T7 V38 hia (33)* gene cassettes, respectively, the *E. coli cer* gene, and a kanamycin resistance gene.

Plasmid BK-96-2-11, that contains a *T7 V38 hia (11)* gene cassette, was linearized with *Bgl* II and dephosphorylated. The *Bgl* II-*BamH* I *T7 V38 hia (11)* gene cassette from DS-2186-2-1 was ligated into BK-96-2-11, generating plasmid DS-2447-2 that contains tandem copies of the *T7 V38 hia (11)* gene in the same orientation (Fig. 9A).

Plasmid DS-2340-2-3 was digested with *EcoR* I and the *T7 V38 hia (33)* gene cassette was subcloned into pUC-BgXb that had been digested with *EcoR* I and dephosphorylated. The resultant plasmid, DS-2440-2 was digested with *Bgl* II and *BamH* I to release the *T7 V38 hia (33)* cassette that was ligated with DS-2340-2-3 that had been linearized with *Bgl* II and dephosphorylated. Plasmid DS-2448-17 contains tandem *T7 V38 hia(33)* genes in the same orientation (Fig. 9B).

Example 9

This Example illustrates the expression of full-length and truncated recombinant *hia* genes.

DNA from expression plasmids prepared as described in the preceding Examples, was introduced into electrocompetent *E. coli* BL21 (DE3) cells using a BioRad electroporator. Cells were grown at 37°C in NZCYM medium using the appropriate antibiotic selection to A_{578} of 0.3 before the addition of lactose to 1.0% for 4 hours. Samples were adjusted to 0.2 OD/µl with SDS-PAGE lysis + loading buffer and the same amount of each protein sample was loaded onto SDS-PAGE gels (ref. 22). Figure 10 illustrates the relative production of rHia (11) proteins from various constructs. As seen in panel A, there is an increase in production with decreased size of rHia. V38- (lane 5) and N52-truncated rHia (lane 6) have significantly higher expression levels than their longer counterparts (lanes 2, 3, 4). In addition, panel B demonstrates that the production of V38 rHia is apparently increased in the presence of the *cer* gene.

Example 10

This Example illustrates the purification of rHia proteins.

All the recombinant Hia proteins were expressed as inclusion bodies in *E. coli* and were purified by the same procedure (Fig.11). *E. coli* cell pellets from 500 ml culture were resuspended in 50 ml of 50 mM Tris-HCl, pH 8.0, containing 0.1 M NaCl, and disrupted by sonication. The extract was centrifuged at 20,000 *g* for 30 min and the resultant supernatant was discarded.

The pellet (PPT_1) was further extracted, in 50 ml of 50 mM Tris-HCl, pH 8.0 containing 0.5% Triton X-100 and 10 mM EDTA, then centrifuged at 20,000 *g* for 30 min, and the supernatant was discarded. The pellet (PPT_2) was further extracted in 50 ml of 50 mM Tris-HCl, pH 8.0, containing 1% octylglucoside, then centrifuged at 20,000 *g* for 30 min, and the supernatant was discarded.

The resultant pellet (PPT_3) obtained after the above extractions contains the inclusion bodies. The pellet was solubilized in 6 ml of 50 mM Tris-HCl, pH 8.0, containing 6 M guanidine and 5 mM DTT. Twelve ml of 50 mM Tris-HCl, pH 8.0 was added to this solution and the mixture was centrifuged at 20,000 *g* for 30 min. The supernatant (SUP_4) was precipitated with polyethylene glycol (PEG) 4000 at a final concentration of 7%. The resultant pellet (PPT_5) was removed by centrifugation at 20,000 *g* for 30 min and the supernatant was precipitated by $(NH_4)_2SO_4$ at 50% saturation. The $(NH_4)_2SO_4$ precipitate was collected by centrifugation at 20,000 *g* for 30 min. The resultant pellet (PPT_6) was dissolved in 2 ml of 50 mM Tris-HCl, pH 8.0, containing 6 M guanidine HCl and 5 mM DTT and the clear solution was purified on a Superdex 200 gel filtration column equilibrated in 50 mM Tris-HCl, pH 8.0, containing 2 M guanidine HCl. The fractions were analysed by SDS-PAGE and those containing the purified rHia were pooled and dialysed overnight at 4°C against PBS, then centrifuged at 20,000 *g* for 30 min. The protein remained soluble under these conditions and glycerol was added to the rHia preparation at a final concentration of 20% for storage at -20°C. SDS-PAGE

analysis of purified V38 rHia (11) and V38 rHia (33) is illustrated in Figure 12. The average yield of the purified V38 rHia proteins is about 10 mg L^{-1} culture.

In order to study the stability of rHia, the purified V38 rHia (11) protein was stored at 4°C with or without glycerol and at -20°C with glycerol. The protein was found to be stable under all three conditions and remained intact for at least eight weeks with repeated freezing and thawing (Fig. 13).

Example 11

This Example illustrates the immunogenicity of V38 rHia (11) and V38 rHia (33) proteins.

Hyperimmune antisera against rHia proteins were produced by immunizing two guinea pigs (Charles River) intramuscularly (i.m.) with 5 µg doses of antigen emulsified in complete Freund's adjuvant (CFA, Difco) on day 1. Animals were boosted on days 14 and 28 with 5 µg doses of protein in incomplete Freund's adjuvant (IFA) and sera were collected on day 42. Anti-Hib strain MinnA and anti- *Haemophilus* type a strain ATCC 9006 antisera were generated using the same protocol, except that a heat-inactivated bacterial preparation was used as the immunogen ($1x10^8$ cfu per dose).

To study the immunogenicity of the V38 rHia proteins, groups of five CD-1 mice (Charles River, Quebec) were immunized s.c. on days 1 and 28 with 0.3, 1, 3, and 10 µg of antigen, in the presence of $AlPO_4$ (alum) (1.5 mg per dose). Blood samples were collected on days 1, 28 and 42. Mice generated significant anti-V38 rHia antibody responses even with a single injection of 0.3 µg antigen (Fig. 14, panel A),

suggesting that both proteins had retained immunogenicity after inclusion body extraction and solubilization. No statistically significant difference was found in the antibody titers induced by the V38 rHia proteins derived from strains 11 or 33.

To study the immunogenicity of the V38 rHia (11) protein in BALB/c mice, groups of five animals (Charles River, Quebec) were immunized s.c. on days 1, 28 and 42 with 0.3, 1, 3, and 10 µg of antigen, in the presence of $AlPO_4$ (1.5 mg per dose). Blood samples were collected on days 1, 14, 28, 42 and 56. High antibody titers were observed in all groups, indicating that the protein is very immunogenic even at 0.3 µg per dose (Fig. 15, panel A).

To study the immunogenicity of the V38 rHia (11) protein in guinea pigs, groups of five animals (Charles River, Quebec) were immunized s.c. on days 1, 28 and 42 with 0.3, 1, 3, and 10 µg of antigen, in the presence of $AlPO_4$ (1.5 mg per dose). Blood samples were collected on days 1, 14, 28, 42 and 56. High antibody titers were observed in all groups, indicating that the protein is also very immunogenic in guinea pigs (Fig. 15, panel B).

Example 12

This Example illustrates the analysis of the protection afforded by anti-rHia antibodies in passive infant rat models of bacteremia.

Pregnant Wistar rats were purchased from Charles River. In the *H. influenzae* type b bacteremia model, groups of 6 to 10 five-day old infant rats were injected s.c. in the dorsal region with 0.1 ml of

guinea pig anti-rHia or anti-strain MinnA antiserum. The control animals received injections with pre-immune sera only. Twenty hours later, the animals were challenged intraperitoneally (i.p.) with 200 to 240 colony-forming units (cfu) of freshly grown Hib strain MinnA (0.1 ml). Blood samples were collected 20 h post-challenge, via cardiac puncture under isoflurane anesthesia and plated on chocolate agar plates. Colonies were counted after one day and the results were statistically analyzed by Fisher's Exact test.

In the *H. influenzae* type a bacteremia model (ref. 23), groups of 9 to 10 five-day old infant rats were injected s.c. in the dorsal region with 0.1 ml of guinea pig anti-rHia or anti-strain ATCC 9006 antiserum. The animals in the control group were injected with guinea pig pre-immune serum. Twenty hours later, the animals were challenged i.p. with 100,000 cfu of freshly grown *H. influenzae* type a strain ATCC 9006 (0.1 ml). Blood samples were collected 20 h post-challenge and analysed as described above.

As shown in Tables 1 and 2 below, the infant rats that were passively immunized with either guinea pig anti-rHia (11) or anti-V38 rHia (11) antisera, were all significantly protected against type a or type b *H. influenzae* caused bacteremia. These results demonstrate that antibodies raised to the slightly truncated Hia protein (V38 rHia) are as efficacious as those raised to the full-length protein at protecting animals against bacteremia caused by type a or type b *H. influenzae*. Such protection afforded by an NTHi-derived recombinant protein against invasive disease

caused by encapsulated bacteria, illustrates the utility of the rHia proteins as vaccine antigens.

Example 13

This Example illustrates the protection afforded by immunization with V38 rHia protein in a chinchilla model of nasopharyngeal colonization.

A nasopharyngeal colonization model has been described by Yang et al (ref. 20). The model works well for those NTHi strains that produce the HMW adhesins, but reproducible colonization could not be established with Hia-producing strains under the same conditions. Repeated attempts to colonize with the prototype Hia-producing NTHi strain 11, were unsuccessful. Colonization was achieved with NTHi strain 33 at 5 x 10^8 cfu per inoculum, compared with only 10^8 cfu required for the prototype HMW-producing NTHi strain 12. Under these conditions, partial protection was observed in animals immunized with 100 µg of V38 rHia (33) and challenged with the homologous NTHi strain 33.

Example 14

This Example illustrates the cloning and sequence analysis of additional *hia* genes from *H. influenzae* strains.

Oligonucleotides (5040.SL and 5039.SL) for PCR amplification were designed based upon the conserved promoter, N-terminal and C-terminal sequences of the *hia* and *hsf* genes and proteins (Fig. 17). The strains chosen for PCR amplification were chosen based upon their reactivity with anti-rHia (11) antisera.

Chromosomal DNA was prepared from NTHi strains 12, 29, 32, M4071, K9 and, K22 and *Haemophilus* type c strain API. PCR amplification was performed as follows: each reaction mixture contained 5 to 100 ng of DNA, 1 µg of each primer, 5 units of taq+ or tsg+ (Sangon) or taq plus long (Stratagene), 2 mM dNTPs, 20 mM Tris-HCl (pH 8.8), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, BSA. Cycling conditions were: 95°C for 1 min, followed by 25 cycles of 95°C for 30 sec, 45°C for 1 min, 72°C for 2 min; then 72°C for 10 min.

The nucleotide and deduced amino acid sequences of the *hia* gene from strain 33 are shown in Figure 18. The predicted Hia protein from strain 33 has a molecular weight of 103.6 kDa and a pI of 9.47. The nucleotide and deduced amino acid sequences of the *hia* gene from strain 32 are shown in Figure 19. The predicted Hia protein from strain 32 has a molecular weight of 70.4 kDa and a pI of 5.67. There is a KDEL sequence present between residues 493 and 496. Such sequences have been associated with anchoring proteins to the endoplasmic reticulum. The deduced strain 32 Hia protein is significantly smaller and has a significantly different pI, however it does contain many of the motifs present in other Hia molecules.

The nucleotide and deduced amino acid sequences of the *hia* gene from strain 29 are shown in Figure 20. The predicted Hia protein from strain 29 has a molecular weight of 114.4 kDa and a pI of 7.58. The nucleotide and deduced amino acid sequences of the *hia* gene from strain K22 are shown in Figure 23. The predicted Hia protein from strain K22 has a molecular weight of 114.4

kDa and a pI of 7.58. The deduced Hia sequences from NTHi strains 29 and K22 were found to be identical. Strain 29 was isolated from a 7-month old child with otitis media in Cleveland, Ohio, while strain K22 was isolated from an aborigine near Kimberly, Australia.

The nucleotide and deduced amino acid sequences of the *hia* gene from strain 4071 are shown in Figure 21. The predicted Hia protein from strain M4071 has a molecular weight of 103.4 kDa and a pI of 9.49. There is a KDEL sequence present between residues 534 and 537.

The nucleotide and deduced amino acid sequences of the *hia* gene from strain K9 are shown in Figure 22. The predicted Hia protein from K9 has a molecular weight of 113.8 kDa and a pI of 6.45.

The nucleotide and deduced amino acid sequences of the *hia* gene from strain type c *Haemophilus* API are shown in Figure 24. The predicted Hia protein from API has a molecular weight of 249.4 kDa and a pI of 5.34. The deduced Hia/Hsf sequence from the type c strain API is nearly identical to the published type b Hsf sequence except for a 60 residue insert. Since the NTHi-based Hia protein provided herein protects in passive models of type a and type b infection, it is likely that it will also protect against type c disease due to sequence similarity between the type b and type c proteins.

The nucleotide and deduced amino acid sequences of the *hia* locus from strain 12 are shown in Figure 25. NTHi strain 12 does not produce Hia. However, part of the *hia* gene can be PCR amplified, there is

inconsistent positive reactivity of SB12 cell lysates with anti-rHia antibody, and there is reactivity with a DNA probe derived from the 3'-end of the strain 11 *hia* gene, on Southern blots. Analysis of the PCR amplified DNA, revealed a 1.8 kb fragment that contains 1 kb of the 3'-end of the upstream HI1732-related gene and 0.8 kb of the 3'-end of the *hia* gene.

PCR amplification using primers that would amplify across the putative junction of these two genes in strain 12, confirmed the genetic composition of the locus. Thus it would appear that strain 12 does not produce Hia because it has suffered a deletion of the 5'-end of the *hia* gene. Figure 27 shows a sequence comparison between the upstream orf of strain 12 and the Rd genome deduced HI1733 protein. Over the region of homology, the two proteins are 95% identical.

An alignment of the deduced Hia sequences from NTHi strains 33, 32, 29, K22, M4071, 11 and K9 and type c strain API compared with *H. influenzae* type b Hsf, the aidA-like (Hsf/Hia) HI1732 gene from the Rd genome, and the *M. catarrhalis* 200 kDa protein from strains 4223 and LES-1 is shown in Figure 28. There is a frame shift in the Rd genome sequence resulting in premature truncation of the HI1732 protein. Additional downstream sequence related to hia, is included here. The asterisks below the sequence indicate conserved residues. The N-terminal (approximately 50 residues) and C-terminal sequences (approximately 150 residues) are highly conserved amongst the *Haemophilus* strains, while some similarity is evident with the *M. catarrhalis* counterpart. Sequence analysis reveals that

there are two potential gene families of Hia proteins, one related to the prototype strain 11 and the other more closely related to strain 33. The strains 11 and K9 proteins appear to be more like the Hsf proteins from the type b, type c or type d *Haemophilus* strains while the strains 33, 32, 29, K22 and M4071 proteins appear to form a second family.

Example 15

This Example describes the construction of plasmid JB-2930-3 that contains a *T7 hia* gene cassette with a truncated S44 strain 11 *hia* gene, the *E. coli cer* gene, and the kanamycin antibiotic resistance gene, and expression of S44 Hia proteins.

PCR primers were designed to amplify the S44 Hia N-terminus of the NTHi strain 11 *hia* gene from the S44 amino acid to an internal *Sty* I site (Fig 29). An *Nde* I site was added at the 5'-end for cloning purposes and the fragment was amplified using plasmid DS-2242-1 as a template. The fragment was cloned into pCR II generating plasmid JB-2910-1-1. The construction scheme is shown in Figure 30. Plasmid JB-2910-1-1 was digested with *Nde* I and Sty I and the 5' PCR *hia* fragment isolated. Plasmid IA-46-5 containing the V38 hia gene was digested with *Nde* I and Sty I and the larger approximately 8.5 kb fragment purified. The two purified fragments were ligated together to produce plasmid JB-2917-1. This plasmid was then digested with *Nde* I and treated with calf intestinal phosphatase (CAP), and into it was cloned the T7 promoter from plasmid IA-46-5. The promoter was cut out using *Nde* I digestion of IA-46-5. The resulting plasmid, JB-2925-3,

was digested with *Bgl* II and *Bam* HI and the *hia* gene was isolated. This fragment was ligated into the *Bgl* II/CAP-treated plasmid BK-2-1-2 to produce plasmid JB-2930-3. This plasmid contains the T7 promoter S44 *hia* gene and *E. coli cer* gene and kanamycin resistance.

The recombinant S44 *hia* vector was transformed into *E. coli* BL21(DE3) for expression studies. The procedure for expression in *E. coli* was as described in Example 9. Figure 31 SDS-PAGE analysis of shows the expression of recombinant S44 *hia* from two different vectors, JB-2930-3 (described above) and pET vector IA-191-3-1. Plasmid IA-191-3-1 is identical to JB-2930-3 except it is a pET vector containing the $lacI^q$ repressor and, therefore, the amount of S44 *Hia* produced is less than the T7 S44 from JB-2930-3. The plasmid is shown, along with plasmid JB-2930-3, Figure 32. Figure 31 shows the S44 *Hia* as a doublet band (lane 3) at approximately 116 kDa. Upon further analysis using purified S44 *hia* from JB-2930-3, the lower band of the doublet was found to have a C-terminal truncation of 94 amino acids, while retaining the expected N-terminus. The purification process used for isolation of the truncated Hia was as described in Example 10.

SUMMARY OF THE DISCLOSURE

In summary of this disclosure, the present invention provides novel isolated and purified nucleic acid molecules encoding full-length and N-terminal truncated *Haemophilus influenzae* adhesin (Hia) proteins from *Haemophilus* which enable protective Hia proteins

to be produced recombinantly. Modifications are possible within the scope of this invention.

REFERENCES

1. Barbour, M.L., R.T. Mayon-White, C. Coles, D.W.M. Crook, and E.R. Moxon. 1995. The impact of conjugate vaccine on carriage of *Haemophilus influenzae* type b. J. Infect. Dis. **171:**93-98.

2. Berkowitz et al. 1987. J. Pediatr. **110:**509.

3. Claesson et al. 1989. J. Pediatr. **114:**97.

4. Black, S.B., H.R. Shinefield, B. Fireman, R. Hiatt, M. Polen, E. Vittinghoff, The Northern California Kaiser Permanent Vaccine Study Center Pediatrics Group. Efficacy in infancy of oligosaccharide conjugate Haemophilus influenzae type b (HbOC) vaccine in a United States population of 61,080 children. 1991. Pediatr. Infect. Dis. J. **10:**97-104.

5. Nitta, D.M., M.A. Jackson, V.F. Burry, and L.C. Olson. 1995. Invasive *Haemophilus influenzae* type f disease. Pediatr. Infect. Dis. J. **14:**157-160.

6. Waggoner-Fountain, L.A., J.O. Hendley, E.J. Cody, V.A. Perriello, and L.G. Donowitz. 1995. The emergence of *Haemophilus influenzae* types e and f as significant pathogens. Clin. Infect. Dis. **21:**1322-1324.

7. Madore, D.V. 1996. Impact of immunization on *Haemophilus influenzae* type b disease. Infectious Agents and Disease **5:**8-20.

8. Bluestone, C.D. 1982. Current concepts in otolaryngology. Otitis media in children: to treat or not to treat? N. Engl. J. Med. **306:**1399-1404.

9. Barenkamp, S.J., and E. Leininger. 1992. Cloning, expression, and DNA sequence analysis of genes encoding nontypeable *Haemophilus influenzae* high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of *Bordetella pertussis*. Infect. Immun. **60:**1302-1313.

10. St. Geme III, J.W., S. Falkow, and S.J. Barenkamp. 1993. High-molecular-weight proteins of nontypeable *Haemophilus influenzae* mediate attachment to human epithelial cells. Proc. Natl. Acad. Sci. USA **90:**2875-2879.

11. Barenkamp, S.J. 1996. Immunization with high-molecular-weight adhesion proteins of nontypeable *Haemophilus influenzae* modifies experimental otitis media in chinchillas. Infect. Immun. **64:**1246-1251.

12. St. Geme, J.W. and D. Cutter. 1995. Evidence that surface fibrils expressed by *Haemophilus influenzae* type b promote attachment to human epithelial cells. Molec. Microbiol. **15:**77-85.

13. Barenkamp, S.J. and J.W. St. Geme. 1996. Identification of a second family of high-molecular-weight adhesion proteins expressed by non-typable *Haemophilus influenzae*. Molec. Microbiol. **19:**1215-1223.

14. St. Geme, J.W., D. Cutter, and S.J. Barenkamp. 1996. Characterization of the genetic locus encoding *Haemophilus influenzae* type b surface fibrils. J. Bact. **178:**6281-6287.

15. Patient, M.E., and D.K. Summers. 1993. ColE1 multimer formation triggers inhibition of *Escherichia coli* cell division. Molec. Microbiol. **9:**1089-1095.

16. O'Hagan, DT. 1992. Oral delivery of vaccines. Formulation and clinical pharmaco kinetic considerations. Clin. Pharmacokinet 22(t): 1-10.

17. Ulmer et al. 1993. Curr. Opinion Invest. Drugs 2:983-989.

18. Lockhoff, O., 1991. Glycolipids as immunomodulators: Synthesis and properties.

19. Nixon-George A., et al., 1990. The adjuvant effect of stearyl tyrosine on a recombinant subunit

hepatitis B surface antigen. J. Immunol 144 (12):4798-4802.

20. Yang, Y-P., S.M. Loosmore, B.J. Underdown, and M.H. Klein. 1998. Nasopharyngeal colonization with nontypeable *Haemophilus influenzae* in chinchillas. Infect. Immun. **66**:1973-1980.

21. Tabor, S., and C.C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074-1078.

22. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. Nature **227**:680-685.

23. Loosmore, S.M., Y-P. Yang, D.C. Coleman, J.M. Shortreed, D.M. England, and M.H. Klein. 1997. Outer membrane protein D15 is conserved among *Haemophilus influenzae* species and may represent a universal protective antigen against invasive disease. Infect. Immun. **65**:3701-3707.

24. Needleman, S.B. and Wunsch, C.D. 1970, J. Mol. Biol. 48:443-453.

25. Sellers, P.H. 1974 On the theory and computation of evolutionary distances. J. Appl. Math(Siam) 26:787-793.

26. Waterman, M.S., Smith, T.F. and Beyer, W.A. 1976. Advan. Math. 20:367-387.

27. Smith, T.F. and Waterman, M.S. 1981 Identification of common molecular subsequences. J. Mol. Biol. 147:195-197.

28. Sobel, E. and Martinez, H.M. 1985 A multiple Sequence Alignment Program. Nucleic Acid Res. 14:363-374.

TABLE 1

Protective effect of guinea pig anti-rHia (full-length) antiserum against type a or b *H. influenzae* in the infant rat model of bacteremia

Group (#)	Guinea pig serum	Anti-rHia antibody titers	No. bacteremic/ No. challenged	Mean cfu/ 100 µl blood
1	Anti-type a	nd	0/10*	0**
2	Anti-rHia	204,800	1/10*	0**
3	Preimmune	<100	7/10	88
Group (#)	Guinea pig serum	anti-rHia antibody titers	No. bacteremic/ No. challenged	Mean cfu/ 2.5 µl blood
4	Anti-MinnA	nd	0/10*	0**
5	Anti-rHia	204,800	1/10*	2**
6	Preimmune	<100	10/10	600

Five-day old infant rats were passively immunized s.c. with 0.1 ml of indicated guinea pig antiserum or preimmune serum. Twenty hours later, infant rats were challenged i.p. with either freshly grown *H. influenzae* type a strain ATCC 9006 (10^5 cfu, 0.1 ml) for groups #1 to 3; or with freshly grown Hib strain MinnA (240 cfu, 0.1 ml) for groups # 4 to 6. Infected animals are defined as >20 cfu recovered from 100 µl of blood for groups #1 to 3; or >30 cfu recovered from 2.5 µl of blood for groups # 4 to 6.

* Fisher exact test. Statistical significance compared to animals in group 3 or 6 was found ($P<0.05$).

** Student's unpaired *t* test. Statistical significance compared to animals in group 3 or 6 was found ($P<0.05$).

nd: not determined.

TABLE 2

Protective effect of guinea pig anti-V38 rHia (SB11) antiserum against type a or b *H. influenzae* in the infant rat model of bacteremia

Group (#)	Guinea pig serum	Anti-rHia antibody titers	No. bacteremic/ No. challenged	Mean cfu/ 20 µl blood
1	Anti-type a	nd	0/6*	0**
2	Anti-rHia	204,800	1/9*	5**
3	Preimmune	<100	5/8	165
Group (#)	Guinea pig serum	anti-rHia antibody titers	No. bacteremic/ No. challenged	Mean cfu/ 2 µl blood
4	Anti-MinnA	nd	0/6*	0**
5	Anti-rHia	204,800	1/9*	2**
6	Preimmune	<100	10/10	820

Five-day old infant rats were passively immunized s.c. with 0.1 ml of indicated guinea pig antiserum or preimmune serum. Twenty hours later, infant rats were challenged i.p. with either freshly grown *H. influenzae* type a strain ATCC 9006 (10^5 cfu, 0.1 ml) for groups #1 to 3; or with freshly grown Hib strain MinnA (190 cfu, 0.1 ml) for groups #4 to 6. Infected animals is defined as >20 cfu recovered from 20 µl of blood for groups #1 to 3; or >30 cfu recovered from 2 µl of blood for groups #4 to 6.

* Fisher exact test. Statistical significance compared to animals in group 3 or 6 was found ($P<0.05$)

** Student's unpaired *t* test. Statistical significance compared to animals in group 3 or 6 was found ($P<0.05$).

nd: Not determined.